

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:16:33 ; Search time 91.75 Seconds

(without alignments)
234.418 Million cell updates/sec

Title: US-09-331-631A-3

Perfect score: 3532
Sequence: 1 MAINTSNLCSLFLFLSLFL.....SSRSRKQKQPLVSLIDPVGF 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1092	30.9	605	1	VCLA_GOSHI
2	1083.5	30.7	588	1	VCLB_GOSHI
3	899	25.5	573	1	GLBI_MAIZE
4	873.5	24.7	614	1	AH11_ARAHY
5	872.5	24.7	639	1	GLCX_SOYBN
6	865.5	24.5	626	1	AH12_ARAHY
7	855.5	24.2	605	1	GLCA_SOYBN
8	832.5	23.6	571	1	CVCA_PEA
9	798	22.6	524	1	SBP_SOYBN
10	760.5	21.5	459	1	VCLC_PEA
11	748	21.2	439	1	GLCB_SOYBN
12	711.5	20.1	463	1	VCL_VICFA
13	704.5	19.9	410	1	VCLB_PEA
14	689.5	19.5	445	1	CANA_CANEN
15	682.5	19.3	445	1	CANA_CANGL
16	568	16.1	436	1	PHSA_PHAVU
17	557	15.8	421	1	PHSB_PHAVU
18	499.5	14.1	386	1	CVCB_PEA
19	414	11.7	275	1	VCLA_PEA
20	226	6.4	1898	1	TRHY_HUMAN
21	222	6.3	1407	1	TRHY_RABIT
22	203	5.7	1549	1	TRHY_SHEEP
23	190.5	5.4	124	1	VCLJ_PEA
24	185	5.2	499	1	GLU2_ORYSA
25	180.5	5.1	500	1	GLU5_ORYSA
26	180.5	5.1	544	1	INVO_AOTTR
27	180.5	5.1	1023	1	GLU4_DROME
28	178	5.0	499	1	GLUB_ORYSA
29	178	5.0	1391	1	MST2_DROXY
30	177.5	5.0	471	1	RU17_XENLA
31	175	5.0	1085	1	YAFB_SCHPO
32	174.5	4.9	987	1	YLX8_CABEL
33	171	4.8	255	1	LP61_ETMTE

34	171	4.8	499	1	GLU4_ORYSA	P14323	oryza sativ
35	168.5	4.8	338	1	LEGB_PEA	P14594	pisum sativ
36	166	4.7	518	1	SSG2_AVEA	P14812	avena sativ
37	165	4.7	495	1	GU12_ORYSA	P07729	oryza sativ
38	165	4.7	499	1	GU11_ORYSA	P07728	oryza sativ
39	165	4.7	518	1	SSG1_AVEA	P12615	avena sativ
40	164	4.6	496	1	GLU2_ORYSA	Q02897	oryza sativ
41	163.5	4.6	793	1	CALD_HUMAN	Q05682	homo sapien
42	163.5	4.6	1240	1	YNI1_YEAST	P53935	saccharomyc
43	163	4.6	516	1	LEGB_GOSHI	P09800	gossypium h
44	163	4.6	708	1	GRF_DICDI	P36417	dicycostell
45	161	4.6	585	1	INVO_HUMAN	P07476	homo sapien

ALIGNMENTS

```

RESULT: 1
VCLA_GOSHI          STANDARD;          PRT;          605 AA.
AC      P09799;
DT      01-MAR-1989 (Rel. 10, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      VICILIN GC72-A PRECURSOR (ALPHA-GLOBULIN A).
OS      Gossypium hirsutum (Upland cotton).
OC      Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC      Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC      Malvales; Malvaceae; Gossypium.
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Chlan C.A., Borroto K., Kamalay J.A., Dure J. III;
RT      "Developmental biochemistry of cottonseed embryogenesis and
RT      germination. XIX. Sequences and genomic organization of the alpha
RT      globulin (vicilin) genes of cottonseed.";
RL      Plant Mol. Biol. 9:335-346(1987).
CC      -1- FUNCTION: SEED STORAGE PROTEIN.
CC      -1- SUBCELLULAR LOCATION: CYTOSOL;ER;MEMBRANE-BOUND VACUOLAR PROTEIN
CC      BODIES.
CC      -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,
CC      CONVICTILIN, CONGLACININ, ETC.).
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL: M19378; AAA33069.1; -.
CC      DR      PIR: S06398; S06398.
CC      DR      HSSP: P50477; ICAX.
CC      DR      INTERPRO: IPR001113; -.
CC      DR      PPM: PF00546; Seedstore_7s; 1.
CC      KW      Seed storage protein; Signal.
CC      FT      SIGNAL          1      23
CC      FT      CHAIN          1      605      VICILIN GC72-A.
CC      SEQUENCE          605 AA; 71049 MW; C9DB9371C976953B CRC64;

```

Query Match 30.9%; Score 1092; DB 1; Length 605;
Best Local Similarity 36.2%; Pred. No. 2.7e-60;
Matches 249; Conservative 118; Mismatches 210; Indels 110; Gaps 17;

```

OY      1 MAINTSNLCSLFLFLSLFLSTVTSGLAE-----SEFD-RQVEYECKRQCGMOLETSQGMK 53
OY      I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
OY      I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
DB      1 MVRKNSVFVFLVFL--SLFLSEGLLCSAKDPGRSRSEDDPOQRYEDCKRRC-QLETRCOTE 57
OY      54 RCVSGQCDKRRFEDIDMSKYNDQDPQDQCCQCRRCRQGESGRQOQYQCRKKEICEEB 113
OY      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      58 O--DKCEDSEYOLKKEQDQDGEDPQRRYQDQCRHOCQOEERRLRP-----HCEQSCRDO 109

```



```

OY 183 EDKKDPQOREVEDCRRCRCEQEPQOYOCQRRCRPOQRHGGGDI INPQRGSGRYEE 242
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 144 EDEQDERQFPFP---RPHQKEERNEEDED--EEQRES-----EE 181
OY 243 GEEKO-----SDMPYFEDERSLSTRTEEGHTSVLENYGRSKLRALKNYRLVLEAN 297
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 182 SESELRHRHKKRPFLEGSNREFLEKNOGRKRLVLORENRSPQLQNLADRILEPNSK 241
OY 298 PNAFLVPTLHLDADAILVTGRCGLAKMIHRDNRESYNLECGDVIRIPAGTFEYLINDNN 357
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 242 PNTLLPNHADADYLLVILNGTALISLVNNDSDYRLQSGDALRVPSTGYTVVVPDNN 301
OY 358 ERHIAKFLQITSTPGQYKEFPAGQONPEPVLSTFSKELLEALNTQALRLGVL----- 413
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 302 ENRLTLTIAIPVKKPGRFSPFLSTEAQSYLQGSNILEASYTKFEIEIKVLEFSRE 361
OY 414 -----GQOR--EGVISAQSOEIRLTDDESRRMIRRGSSRGPNLFKRLYSN 466
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 362 EGOQOGEORLOEYVIEISKEQIRALSKRAKSSSKRT-----SEDKPPLNSRDPYISN 417
OY 467 KYGAQVEKPEYDQLQDMVSVFIANITQSGMGPFFNTRSTKVVVAVSGADVEMACP 526
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 418 KLGKFEITPEKPPQRLDIFLIYDMNGALLPHFNSKALIVILVINGDANIELV-- 475
OY 527 HLSGRHG--RRGKRHEEEDVHYEQVAKRLSKREALIVPVGHVPVYSSGNEMLLRA 584
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 476 -----GKEQOQEOQOGEQPLEVRKRYARELSEODIFVIPAGYVV--NATSNLNFPA 526
OY 585 FGIAONNHNENFLAGRRERNVLOIEPQAMELAFAPARKVEEELNSODESIFPPGRHO 644
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 527 IGINAENNRNFLAGSDQVYISQIPQVOELAFPGSAQAVKRLKRNRSYFVDAQPKK 586
OY 645 QOQSRSTKQOQPLVSL 661
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 587 EEGNGKRK--GPLSSIL 601

RESULT 8
CVCA_PEA STANDARD: PRT: 571 AA.
ID CVCA_PEA STANDARD: PRT: 571 AA.
AC P13915;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE CONVICTILIN PRECURSOR.
GN CVCA.
OS Pisum sativum (Garden pea).
OC Eukaryota: Viridiplantae: Tracheophyta: Spermatophyta:
OC Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids 1;
OC Fabales: Fabaceae: Papilionoideae: Pisum.
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-CV. FELTHAM FIRST.
RX MEDLINE=88326208: PubMed=3415641;
RA Bown D., Ellis T.H.N., Gatehouse J.A.;
RT "The sequence of a gene encoding convicillin from pea (Pisum sativum
RT L.) shows that convicillin differs from vicillin by an insertion near
RT the N-terminus."
RL Biochem. J. 251:717-726(1988).
CC -!- FUNCTION: SEED STORAGE PROTEIN.
CC -!- SUBCELLULAR LOCATION: CYTLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN
CC BODIES.
CC -!- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,
CC CONVICTILIN, CONGLYCININ, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----

```

```

DR EMBL: X06398; CAA29695.1; -.
DR PIR: S00566; S00566.
DR HSSP: P02853; 2PIL.
DR INTERPRO: IPR00113; -.
DR PFAM: PF00546; Seedstore_7s; 1.
KW Seed storage protein; Multigene family; Signal.
FT SIGNAL
FT CHAIN
SQ SEQUENCE 571 AA; 66989 MW; 749CFBEB2D16D57B CRC64;
CONVICTILIN.

Query Match 23.6%; Score 832.5; DB 1: Length 571;
Best Local Similarity 32.7%; Pred. No. 2,4e-44;
Matches 204; Conservative 110; Mismatches 196; Indels 113; Gaps 16;

OY 72 YDNQDDPQTDCCQRCRCEQEPQOYOCQRRCRCEQEEYEVYRQDDPQOYEQOQ 131
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 27 YANYDGCSETRYPGQREORQEG-----EKEKRHGEKRPSTKEEHEE 70
OY 132 RQQRHETPRHMQTCQRCRERYEKERKQOK---RYEQOREDEKEYEERKKEEDNR 187
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 71 EKQKY-----RYQREKQKEQVQPGREMEREREDEQYEEWRGQRE 114
OY 188 DPQOREYEDCRRCRCEQEPQOYOCQRRCRCEQORQGRGDLINPQRGSGRYEEGEO 247
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 115 DPEE-----KALRHRER-----TKDRRH-----QREGEERSSSESO 150
OY 248 SDNPYFEDERSLSTRTEEGHTSVLENYGRSKLRALKNYRLVLEANPAFLVPTLH 307
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 151 HNRPFLEKSNKFLTPENNGHTRLRQDRKSDLEFNQNLRLVEYRAKPHPIFLPQHI 210
OY 308 DADAILVTGRCGLAKMIHRDNRESYNLECGDVIRIPAGTFEYLINDNNERHIAKFLQ 367
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 211 DADLLVLTNGKALITVLSPNDRNSYLERGDTIKIPAGTSTYLVNODDEDLRVVDVI 270
OY 368 TISTPGQYKEFPAGQONPEPVLSTFSKELLEALNTQALRLGVL----- 413
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 271 PVNRGKFEAF--GISENKNQYLRGFSKNILASLTKYETLEKYLEBOEKKPOOLRDR 328
OY 414 -----GQOREGYIISASQOIRLTDDESRRMIRRGSSRGPNLFKRLYSN 468
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 329 KRTQOGEERD-ALIKVSRQIEELKRLAKSS-----KKLPDEFEPFLNSRKPEYSNKF 383
OY 469 GOAVEYKPE-DYRQLODMVSVFIANITQSGMGPFFNTRSTKVVVAVSGADVEMACP 527
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 384 GKLEFETPEKKYQLODDILVSCVEINKGALMLPHYSRAIVLVLTNGKGNLELL 440
OY 528 LSGRHGRGCRKHEEE-----DVHTEQVAKRLSKREALIVPVGHVPVYSSGNEML 580
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 441 -----GKNEQOERDEKRRERNEVQRYEARISPDGVYIIPAGHPVAISASNLML 490
OY 581 LLEAFGINAONHNENFLAGRRERNVLOIEPQAMELAFAPARKVEEELNSODESIFPPG 639
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 491 L--GCIANKKNORNFLSSDQNVISQIENPVELTFFROSSQOEVNRLKRNQSHFASNE 548
OY 640 PROHQQOQSRSTKQOQPLVSL 662
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 549 PEQKEESQR---KRSPSSVLD 568

RESULT 9
SBP_SOYBN STANDARD: PRT: 524 AA.
ID SBP_SOYBN STANDARD: PRT: 524 AA.
AC 004672;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE SUCROSE-BINDING PROTEIN PRECURSOR (SBP).
GN SBP.
OS Glycine max (Soybean).
OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:
OC Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids 1;
OC Fabales: Fabaceae: Papilionoideae: Glycine.

```

[illegible]

```

Db 377 RGH1Q1SCPHMSSSSSHK-----HDKSPSYHRTSSDLKCGMWVYVPGPFTVIASNK 431
QY 578 ENLLFAFGINAQNHNENFLAGRENNYLOQIEPOAMELAFAPRKEVEEL----- 628
Db 432 ENLLMICEVNAARDKKFTFAGKD-NIVSSLDNVAKEELAFNYPSBMVNGVFLDQFLERK 490
OY 629 -----NSODESTFFPGPROHOOOSSS 650
Db 491 LIGRLYHLPHKRDRESEFFFPFLPREGRRA 522

RESULT 10
VCLC_PEA STANDARD: PRT: 459 AA.
AC P13918;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE VICILIN PRECURSOR.
OC Pisum sativum (garden pea).
OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:
OC Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids I:
OC Fabales; Fabaceae; Papilionoideae; Pisum.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEEDLING:
RA Higgins T.J.V., Newbiggin E.J., Spencer D., Llewellyn D.J., Craig S.;
RT "The sequence of a pea vicilin gene and its expression in transgenic
RT tobacco".
RL Plant Mol. Biol. 11:683-695(1988).
[2]
SEQUENCE OF 27-459 FROM N.A. (CLONE PDBU9).
RP STRAIN=CV. FELTHAM FIRST:
RC MEDLINE=88326226; Pubmed=3046604;
RA Watson M.D., Lambert N., Delauney A., Yarwood J.N., Croy R.R.D.,
RA Gatehouse J.A., Wright D.J., Boulter D.;
RT "Isolation and expression of a pea vicilin cDNA in the yeast
RT Saccharomyces cerevisiae".
RL Biochem. J. 251:857-864(1988).
CC -1- FUNCTION: SEED STORAGE PROTEIN.
CC -1- SUBCELLULAR LOCATION: CYTOLEONARY MEMBRANE-BOUND VACUOLAR PROTEIN
CC BODIES.
CC -1- MISCELLANEOUS: THERE ARE AT LEAST 11 GENES FOR VICILIN SUBUNITS.
CC -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASCOLIN, VICILIN,
CC CONVICILIN, CONGLYCININ, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14076; CA32239.1; -
DR EMBL: Y00722; CA68708.1; -
DR PIR: S00567; S00567.
DR PIR: S08505; S08505.
DR HSP: P50477; ICAM.
DR INTERPRO: IPR001113; -
DR PFM: PF00546; Seedstore_7s; 1.
KW seed storage protein; Multigene family; signal.
FT SIGNAL 1 28
FT CHAIN 29 459 VICILIN.
FT CONFLICT 375 375 V -> L (IN REF. 2).
FT CONFLICT 409 409 I -> V (IN REF. 2).
SQ SEQUENCE 459 AA; 52231 MW; E2FA390B95451D29 CRC64;

Query Match 21.5%: Score 760.5; DB 1; Length 459;
Best Local Similarity 37.6%: Pred. NO. 5e-40;
Matches 161; Conservative 79; Mismatches 159; Indels 29; Gaps 5;

```


RX	MEDLINE=88067789; PubMed=3684610.
RA	Bassuener R.,van Nong H., Jung R., Saalbach G., Muentz K.;
RT	"The primary structure of the predominating vicillin storage protein
RL	subunit from field bean seeds (Vicia faba L. var. minor cv. Filbo).";
CC	Nucleic Acids Res. 15:9609-9609(1987).
CC	-1- FUNCTION: SEED STORAGE PROTEIN.
CC	-1- SUBCELLULAR LOCATION: CYTOSOL;NON-MEMBRANE-BOUND VACUOLAR PROTEIN BODIES.
CC	-1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONVICILIN, CONGLYCININ, ETC.).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	-----
DR	EMBL; Y00506; CAA68559.1; -;
DR	EMBL; X00462; CAA68525.1; -;
DR	PIR; S06309; S06309.
DR	PIR; A27288; A27288.
DR	HSSP; P50477; ICAM.
DR	INTERPRO; IPR001113; -;
KW	PFAM; PF00546; Seedstore_7s; 1.
KW	Seed storage protein; Signal.
FT SIGNAL	1 27
FT CHAIN	1 463
FT VICILIN.	
FT CONFLICT	257 257 R -> K (IN REF. 2).
FT CONFLICT	443 443 R -> Q (IN REF. 2).
QO	SEQUENCE 463 AA; 52694 MW; SCDD9EFDE0D5E6B5 CRC64;

Query Match	20.1%;	Score 711.5;	DB 1;	Length 463;
Best Local Similarity	35.8%;	Pred. No. 5.2e-37;		
Matches 157;	Conservative 86;	Mismatches 159;	Indels 37;	Gaps 6;

QY	249	DNPYFEPERSLSTFEETFEEGHISVLENFYGRSKLLTALKYVRLVLEANNAPVLPT	HD 308
Db	33	DNPVFEASNRFQTLFENBNGHILLOKFDHSHKLEMLQVRLLEYKSKRHTTFLFOOTD	92
QY	309	ADAILLVTRGALKMTIHDRNRESYNLECGDVIRIPAGTFTVYLINDRNNERLHIAFLQT	368
Db	93	ADFLVLVLSKAILTVLLPDRNRSFSLERDGTIKLPAGTIGYLVNRDEEDLRYLDVIP	152
QY	369	ISTPGYQKKEFPFAGGQNPBEYLTSTFSKTELEALNQAELRLVLOQ	415
Db	153	VNRGEFQSTFLSGNONQPSILSGFSKNITLEASFNTDYKTEIKVLELLEHGKERYHRRGLK	212
QY	416	-----OREVIIISASQEQIRELTRDSDSESRMHIRRGSESSRGPNLFPKRPILSNKYG	469
Db	213	DRQROGEENVYIKRSKQJLELNKAKSS-----KKSTSESEFPNLRREPIYNSKKG	268
QY	470	QAEVKEKEDYROLQDMDVSVFIANTITQGSMMGPENTRSTKVVVVVSGADVEMACPHLS	529
Db	269	KFEETIKRPNQDLODLNIFNVYEINEGSLLRPHYSRAIVITVNEKGDFELVQGRNE	328
QY	530	GRHGGRGKGRHREEDVHTEOV-----KARLSKREALIVVGVHPYVSVSGNENLL	582
Db	329	NQOGLR--EYDEDEKQGEDELRKOVYNTKAKLSPDVAIVAGYVVALKASSNINLV--	384
QY	583	FAPFIGNONNHENFLAGRERNVLOQIEPQAMELAFAPKREVEELNQSDESTFEFPGRO	642
Db	385	-GFCINANNQRYFLAGEEDNVISQIHKPYELAFESQAGEVDTLLENQKSHFANNAQPR	443
QY	643	HQOQSSRSTKOQDPVLSL 661	
Db	444	EREGSQEIKDH--LYSIL 460	

RESULT 13
VCLB_PEA

ID	VCID	PEIN	STANDARD;	PRT;	410 AA.
AC	AC	21-JUL-1986	(Rel. 01, Created)		
DT	DT	21-JUL-1986	(Rel. 01, Last sequence update)		
DT	DT	01-JAN-1990	(Rel. 13, Last annotation update)		
DE	PROVICILIN	PRECUSOR	(TYPE B (FRAGMENT)).		
OS	Pisum sativum	(Garden pea).			
OC	Eukaryota:	Viridiplantae:	Embryophyta:	Tracheophyta:	Spermatophyta:
OC	Magnoliophyta:	endicots/ledons:	core eudicots:	Rosales:	euroside 1:
OC	Fagales:	Fabaceae:	Papilionoideae:	Pisum.	
FN	[1]				
RP	SEQUENCE	FROM N.A. (CLONES PDUB7 AND PDUB4).			
RC	STRAIN=CV.	FELTHAM FIRST.			
RX	MEDLINE=83220791;	PubMed=6667941;			
RA	Lycett G.W., Delauney A.J., Gatehouse J.A., Gilroy J., Croy R.R.D.,				
RA	Boulter D.;				
RT	"The vicilin gene family of pea (Pisum sativum L.): a complete cDNA				
RT	coding sequence for preprovicilin."				
RL	Nucleic Acids Res. 11:2367-2380(1983).				
CC	-1- FUNCTION: SEED STORAGE PROTEIN.				
CC	-1- SUBCELLULAR LOCATION: CYTOSOL; VACUOLAR MEMBRANE-BOUND VACUOLAR PROTEIN				
CC	BODIES.				
CC	-1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,				
CC	CONVICILIN, CONGLYCININ, ETC.).				
CC	PIR: A03344; EFWMB.				
DR	HSSP: P50477; ICAM.				
DR	INTERPRO: IPR001113; "				
DR	PRAM: PF00546; Seedstore_7s. 1.				
DR	Seed storage protein; Glycoprotein; Multigene family; Signal.				
KW	STGNAL	1	15		
FT	CHAIN	16	>410		
FT	STE	221	222		
FT	CARBOHYD	359	359		
FT	NON_TER	410	410		
Q	SEQUENCE	410 AA:	46385 MW:	8AF6CE8E5A316FA2 CRC64:	

Query Match	19.9%;	Score 704.5;	DB 1;	Length 410;
Best Local Similarity	37.2%;	Pred. No. 1.2e-36;		
Matches 148;	Conservative 76;	Mismatches 145;	Indels 29;	Gaps 5

QY	249	DNYEYPERLSLSTFEEDECHI	SVLENFYRSKILRLAKVNRVLVLEA	NPANNAFLPHLD	308
Db	20	ENPIEFKSNRQYLYLEMBNGH	IRILCKFDKRSKIFEMLOYMRILLET	SKSHYTLFLPYT	79
QY	309	ADAILVTGGRGLKMIHRDRES	YNLECGDVIRIPAGTYFLYLNR	NNERLIIAKFLQT	368
Db	80	ADFLIVYLSKATLYLVKSN	DRNSPFIENLRDAIKLPAGS	IATYAFNRDNEPRLD	139
QY	369	ISTPGVCKEFPFAGGQNP	PEYLTSTFSKELLFALINTD	AEKLRGVLCOR-----	417
Db	140	VNRFGOLSLSTGTONOKSS	LSGFSKNILEAFNTYEEIEK	VLLEQOODEPQHRSLK	199
QY	418	-----EEVYISASOEOL	RELTRODSESRWHIRRGES	SGRPVLFKPKRPLYSKYG	469
Db	200	DKROELNEEWLYKVSQD	LELSKNAKSS-----	KSVSSESGPFLKSRNPYS	KFG 255
QY	470	QAYEVPEDYROLQDMVSY	FIANTITOGSMKGFENFR	STKTVVVVNSGEADVEM	ACPHLS 529
Db	256	KFEFLPEKKQOQLDD	LDIFVNSVDIKKGSLL	LRPNYSRAIVITYVB	GKDFELYQGR-- 313
QY	530	GRHGGRGGRHREEDVY	--HYEOVKARLSKREAI	VYPGVHVPVSSGNEILL	FAFGI 567
Db	314	NENQKGKENDKEE	OEETSQOVOLYRAKLS	PRGQVFIAPGHPALIN	ASSDLNTI--GLGI 371
QY	588	NAONNHENLAGRE	NRVYLOOLEQAMLA	FAARKEVE 625	
Db	372	MAENNERNPLAGE	DVNTISOVERPVKEL	APGSSHEVD 409	

RESULT	14
CANA_CANEN	
ID	CANA_CANEN
STANDARD;	
PRT;	445 AA

```

AC P50477: 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CANAVALIN PRECURSOR.
OS Canavalia ensiformis (Jack bean) (Horse bean) .
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Canavalia.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COFFLEON;
RX NG J.D., Stinchcombe T., Ko T.-P., Alexander E., McPherson A.;
RX "PCR cloning of the full-length cDNA for the seed protein canavalin
RX from the jack bean plant, Canavalia ensiformis.";
RL Plant Mol. Biol. 18:147-149(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=COFFLEON;
RX MEDLINE=94143475; PubMed=8310055;
RX NG J.D., Ko T.-P., McPherson A.;
RX "Cloning, expression, and crystallization of jack bean (Canavalia
RX ensiformis) canavalin.";
RL Plant Physiol. 101:713-728(1993).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=94143476; PubMed=8310056;
RX Ko T.-P., Ng J.D., McPherson A.;
RX "The three-dimensional structure of canavalin from jack bean
RX (Canavalia ensiformis)." ;
RL Plant Physiol. 101:729-744(1993).
CC -!- FUNCTION: SEED STORAGE PROTEIN.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,
CC CONVICILIN, CONGLYCININ, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59467; CAA42075.1; -
DR PDB: ICAU; 3I-OCT-93.
DR PDB: ICAY; 3I-OCT-93.
DR PDB: ICAM; 3I-OCT-93.
DR PDB: ICAX; 3I-OCT-93.
DR INTERPRO: IPR001113; -
DR Pfam: PF00546; Seedstore_7s; 1.
KW Seed storage protein; Signal; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 445 CANAVALIN.
SEQUENCE 445 AA; 50326 MW; 30383C5F83A1E9B7 CRC64;

```

Query Match	19.5%	Score 689.5;	DB 1;	Length 445;
Best Local Similarity	33.6%	Pred. No. 1.1e-35;		
Matches .143;	Conservative 97;	Mismatches 163;	Indels 23;	Gaps 5;
OY	235	GCSCGHEGCEKOSDNPPYDFERSLSTRFRFECHISVLENFYGSKRLAKNRYLL	294	
		: : : : : : : : :		
Db	32	GGEADDEESRAONNPLFRSNKFLTFKQHGSLRLQGFNETELEMRLROYLEY	91	
OY	295	EANPAEVLPHILDAAILVTGGRGALKMTHRDRESYNLECGDVIRIPAGTFEYLNR	354	
		: : : : : : : : : : : : : :		
Db	92	CSKPTLLPRHNSDDLVLVLEGCALIVLVNPDGRDYKIKDQGAIRIKQGTPEYLNP	151	
OY	355	DNNEELIAKFLQITSTPGQKKEPPRAGCONDEPLSTFSKELLEALNATQAEKL-RQVL	413	
		: : : : : : : : : : : : :		
Db	152	DNNOULRIKALITFRRGTEDEFLSTKRLPYLSAFSKNFLEAVSDSYDETEQLL	211	

```

OY 414 GQOGEVITISASOEDIRELTRDSDSESRMIIIRGCGESSRGPIVFNKRPILYSNKYQATE 473
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 212 QEOEGCVIVKPKRQIDGITSKHAOSSSKTIL----SSODKPFNIRSDPIYSNNYGLYE 267
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 474 VKPEEDROLQMDVSVFLANTIGGSSMMGPFFENTRSTVYVVVVAAGSEADVEMACPILSGRIG 533
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 268 ITPENKSOIRLDILILNLOKMEGALLFVPHNYSRAFTILVANEGRAREVELY----- 318
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 544 GRGCKRHEEDEDVHYEDVK---ARKSKRELIVYVGHVYVVSNGENLILFAFGINAQ 550
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 319 ----GLEDOQOQGLSEMLRRYATFLLEGDTIVPSSFPAALKAASDLNV--GIGVNAE 372
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 591 NNHNENFLAGRENVLOQIEPQAMALEAPARKVEYELFNSODESIFPPGROHQOOSRS 650
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 373 NNNERNFLAGHKENVIHQIPRGVSDLTGPGSGSEVEVELLENQKESYFVDOGPHNIDAGKA 432
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 651 TKQOOP 656
   :::|
Db 433 RRHILP 438

```

```

RESULT 15
ID CANA_CANGT STANDARD: PRT: 445 AA.
AC P10562;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CANAVALIN PRECURSOR.
OS Canavalia gladiata (Sword bean) (Japanese jack bean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1;
OC Fabales; Fabaceae; Papilionoideae; Canavalia.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-40.
RC TISSUE=SEED;
RX MEDLINE=88111636; PubMed=3338449;
RY Yamauchi D., Nakamura K., Asahi T., Minamikawa T.;
RT "CNAs for canavalin and concanavalin A from Canavalia gladiata
RT seeds. Nucleotide sequence of cDNA for canavalin and RNA blot
RT analysis of canavalin and concanavalin A mRNAs in developing seeds.";
RL Eur. J. Biochem. 170:515-520(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RX MEDLINE=89296493; PubMed=2740227;
RY Takai Y., Yamauchi D., Minamikawa T.;
RT "Nucleotide sequence of the canavalin gene from Canavalia gladiata
RT seeds.";
RL Nucleic Acids Res. 17:4381-4381(1989).
CC -1- FUNCTION: SEED STORAGE PROTEIN.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,
CC CONVICILIN, CONGLYCININ, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/sib-sib.ch).
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X06733; CAA29910.1; -
DR EMBL: X15076; CAA33172.1; -
DR PIR: S00281; S00281.
DR HSSP: P50477; ICAM.
DR INTERPRO: IPR001113; -
DR Pfam: PF00546; Seedstore_7s; 1.
KW Seed storage protein, Signal.
FT SIGNAL
FT CHAIN 27 445 CANAVALIN.

```

FT CONFLICT 161 161 N -> K (IN REF. 2).
SQ SEQUENCE 445 AA; 50289 MW; 739B59CC2292EA80 CRC64;

Query Match 19.3% Score 682.5; DB 1; Length 445;

Best Local Similarity 33.1% Pred. NO. 3e-35;
Matches 141; Conservative 98; Mismatches 164; Indels 23; Gaps 5;

```
OY 235 GGSGRYEGEGEKOSDNPYFEDERSLSTRFTEEGHISYLENFYGRSKLLRALKNYRLVLL 294
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 32 GGEADEESEEERAAQNNPYLFPSNKKFLTLFKNHGSLRLQRFNEDTEKLENLRDYRVLEY 91
OY 295 EANNAPVLPYTHLDADAILVTGRCALKMTHRDNRESYNLECGDVIRTPAGTFYLLNR 354
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 92 CSKPNLTLLPHHSDDLLVLEGOAILVYNPDGRDPTYKLDQGDALKIOAGTFPYLLNP 151
OY 355 DNNERLHAKFLOTISTPGQYKEFFPAGGONPEPYLSTFSKEILEALNTQAERL-RGYL 413
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 152 DNMGNLILNFATFRRRGTYVEDFLSTKRLPSYLSAFSKNPLEASYDSPYDEIEOTLL 211
OY 414 GOOREGVIIISASQIRELTRDSESRWHIRRGESSRGPNLFNKRPLYSNKYGQAYE 473
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 212 QEEGEGVIVKMPKQDQIEISKHAGSSRKT-----SSQDKPFNLRSDPYISNNYKLYE 267
OY 474 VKPEDYROLQDMDVSVFIANITGSGMMGPFFNTRSTKVVVVASGEADVEMACPRLSGRHG 533
    : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 268 ITPEKNSQLRDLIDLINCLQMNNEGALFVPHYNSRATVILVANEGRAEVELV----- 318
OY 534 GRGGRKRHEEEDVHYEOK---ARLSKREAIYVPVGHVYFVSSGNENLLFAFGINAQ 590
    319 ----GLEOQOQOQLESQMLKRYAATLSEGLVLPSSFPVALKAASDLNMV--GIGVNAE 372
OY 591 NNHENFLAGREERNVLOQIEPQAMELAFAPRKEVEELFNSODESIFFPPGPROHQOQSSRS 650
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 373 NNERNFLAGKKNENYRQIPROVSDLTFPGSGEEVEELLENNKESYFYDGGQPRHIDAGKA 432
OY 651 TKOQOP 656
    : |
Db 433 RRAHLP 438
```

Search completed: March 1, 2001, 16:16:37
Job time: 417 sec

